



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Efficient Checking of Individual Rewards Properties in Markov Population Models

**Citation for published version:**

Bortolussi, L & Hillston, J 2015, Efficient Checking of Individual Rewards Properties in Markov Population Models. in *Proceedings for the 13th Workshop on Quantitative Aspects of Programming Languages*. Workshop on Quantitative Analysis of Programming Languages 2015, London, United Kingdom, 11/04/15. <<http://qapl15.inria.fr/program.html>>

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

Proceedings for the 13th Workshop on Quantitative Aspects of Programming Languages

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Efficient Checking of Individual Rewards Properties in Markov Population Models

Luca Bortolussi

Modelling and Simulation Group, Saarland University, Germany  
DMG, University of Trieste, Italy  
CNR/ISTI, Pisa, Italy  
luca@dmi.units.it

Jane Hillston

School of Informatics, University of Edinburgh  
jane.hillston@ed.ac.uk

In recent years fluid approaches to the analysis of Markov populations models have been demonstrated to have great pragmatic value. Initially developed to estimate the behaviour of the system in terms of the expected values of population counts, the fluid approach has subsequently been extended to more sophisticated interrogations of models through its embedding within model checking procedures. In this paper we extend recent work on checking CSL properties of individual agents within a Markovian population model, to consider the checking of properties which incorporate rewards.

## 1 Introduction

We currently face the scientific and engineering challenge of designing large scale systems, where many autonomous components interact with each other and with humans in an open environment. Examples include power generation and distribution in smart grids, bike and car sharing systems, e-vehicles and public transportation in smart cities. In order to properly design such collective adaptive systems (CAS), mathematical and computational modelling with predictive capabilities is essential. However, the large scale of such systems, and of their corresponding models, exacerbates state space explosion, introducing exceptional computational challenges. In particular, computer-aided verification of formal properties, recognised good practice in software development, protocol design, and so on [15], is hindered by the scalability issues. Moreover the open and uncertain nature of such CAS calls for the use of stochastic models capable of quantifying uncertainty, which introduces further challenges in analysing large scale models [29]. Analysis of stochastic models, in fact, is a computationally intensive procedure. Numerical approaches [29] suffer greatly from state space explosion of models of CAS systems, and statistical methods based on simulation also require a lot of computational effort.

A promising recent approach to modelling systems which consist of interacting populations of entities is based on fluid approximation [13], in which the discrete evolution of the system is replaced by a continuous approximation where the evolution of the state space is captured by a set of ordinary differential equations. Whilst some predictive modelling can be carried out in terms of an approximation in which the entire model is treated in the fluid limit, this approach does not transfer to verification of properties of stochastic systems as all variability is lost in the continuous approximation. Instead, in [11], the authors show how model checking properties of a single entity (or small group of entities) within a collective system can be carried out by retaining the discreteness of the featured entity/entities but placing this in the context of a fluid approximation of the rest of the system. This offered the first possibility to formally verify the properties of extremely large scale systems without the computational cost of statistical model checking. There is a limitation to the properties of a single entity (or small set of entities). Nevertheless, this class of properties is quite common in performance models and in network epidemics

[19]. For instance, in client/server systems, we may be interested in quality-of-service metrics, like the expected service time [25]. In network epidemics, instead, we may be interested in properties connected with the probability of a single node being infected within a certain amount of time, or in the probability of being patched before being infected [21]. Other classes of systems can be naturally queried from the perspective of a single agent, including ecological models [31] (survival chances of an individual or foraging patterns), single enzyme kinetics in biochemistry [28] (performance of an enzyme), but also crowd models [26] or public transportation models in a smart city.

In [11], however, only time bounded properties of classic CSL were considered. In this paper we consider the extension of this approach to extended CSL which supports properties based on rewards. Overlaying a reward structure [20] on a continuous time Markov chain (CTMC) allows more sophisticated reasoning about the behaviour of the system, and CSL with rewards can assess properties related to the accumulated or instantaneous values associated with particular states or transitions of the system.

An alternative approach to fluid model checking has also been investigated by Bortolussi and Lanciani [14]. Their work is based on a second-order fluid approximation known as Linear Noise Approximation [34], and allows them to lift local specification to collective ones. This can be regarded as a functional version of the Central Limit Approximation [23]. Thus the properties that they consider are first expressed as a property of an individual agent, specified by a deterministic timed automaton with a single clock. A further use of mean field approximation in model checking has recently been developed for discrete time, synchronous-clock population processes by Loreti *et al.* [24], although the approach followed in this case is somewhat different as it is an on-the-fly model checker, only examining states as they are required for checking the property.

In this paper we extend the previous work on fluid model checking to incorporate properties of single agents expressed in extended CSL, i.e. properties with cumulative and instantaneous rewards. We also consider reward properties at steady state when such an equilibrium exists in the studied system. This involves adding and evaluating rewards to the inhomogeneous continuous time Markov chain (ICTMC). Reward properties have been considered previously in the context of fluid approximation, but in a collective perspective [33], also considering hybrid approximations caused by the addition of a feedback from the reward to the system [30].

The paper is organised as follows: in Section 2 we introduce the relevant background material, including Markov Population Models, Fluid approximation, CSL, and fluid model checking. In Section 3, we discuss reward properties of individual agents and their fluid approximation. Conclusions are drawn in Section 4.

## 2 Background

### 2.1 Markov Population Models

In this section, we will introduce a simple language to construct Markov models of populations of interacting agents. We will consider models of processes evolving in continuous time, although a similar theory can be considered for discrete-time models (see, for instance, [13, 24]). In principle, we can have different classes of agents, and many agents for each class in the system. Furthermore, the number of agents can change at runtime, due to birth or death events. Models of this kind include computer networks, where each node (e.g. server, client) of the network is an agent [25], biological systems (in which molecules are the agents) [32], and so on. However, to keep notation simple, we will assume here that the number of agents is conserved and equal to  $N$  (making a closed world assumption). Furthermore, in the notation we do not distinguish between different classes of agents.

In particular, let us assume that each agent is a finite state machine, with internal states taken from a finite set  $S = \{1, 2, \dots, n\}$ . We have a population of  $N$  agents, and denote the state of agent  $i$  at time  $t$ , for  $i = 1, \dots, N$ , by  $Y_i^{(N)}(t) \in S$ . Note that we make explicit the dependence on  $N$ , the total population size. A configuration of a system is thus represented by the tuple  $(Y_1^{(N)}, \dots, Y_N^{(N)})$  — each agent is treated as a distinct individual with identity conferred by the position in the vector. However, when dealing with population models, it is customary to assume that single agents in the same internal state cannot be distinguished, hence we can move to the *collective representation* by introducing  $n$  counting variables:

$$X_j^{(N)} = \sum_{i=1}^N \mathbf{1}\{Y_i^{(N)} = j\}. \quad (2.1)$$

Note that the vector  $\mathbf{X}^{(N)} = (X_1^{(N)}, \dots, X_n^{(N)})$  has dimension independent of  $N$ ; this will be referred to as the *collective, population, or counting* vector. The domain of each variable  $X_j^{(N)}$  is  $\{0, \dots, N\}$ , and, by the closed world assumption,  $\sum_{j=1}^n X_j^{(N)} = N$ . Let  $\mathcal{S}^{(N)}$  denote the subset of vectors of  $\{1, \dots, N\}^n$  that satisfy this constraint.

In order to capture the dynamics of such models, we will specify a set of possible events, or transitions, that can change the state of the system. Each such event will involve just a small, fixed, number of agents. Events are stochastic, and take an exponentially distributed time to happen, with a rate depending on the current global state of the system. The set of events, or transitions,  $\mathcal{T}^{(N)}$ , is made up of elements  $\tau \in \mathcal{T}^{(N)}$ , which are pairs  $\tau = (R_\tau, r_\tau^{(N)})$ . More specifically,  $R_\tau$  is a *multi-set of update rules* of the form  $i \rightarrow j$ , specifying that an agent changes state from  $i$  to  $j$  when the event fires. As  $R_\tau$  is a multiset, we can describe events in which two or more agents in state  $i$  synchronise and change state to  $j$ . The exact number of agents synchronising is captured by the multiplicity of rule  $i \rightarrow j$  in  $R_\tau$ ; we denote this by  $m_{\tau, i \rightarrow j}$ . Note that  $R_\tau$  is independent of  $N$ , i.e. each transition involves a finite and fixed number of individuals.

In order to model the effect of event  $\tau$  on the population vector, we will construct from  $R_\tau$  the *update vector*  $\mathbf{v}_\tau$ :  $\mathbf{v}_{\tau, i} = \sum_{(i \rightarrow j) \in R_\tau} m_{\tau, i \rightarrow j} \mathbf{e}_j - \sum_{(i \rightarrow j) \in R_\tau} m_{\tau, i \rightarrow j} \mathbf{e}_i$ , where  $\mathbf{e}_i$  is the vector equal to one in position  $i$  and zero elsewhere. Then, event  $\tau$  changes the state from  $\mathbf{X}^{(N)}$  to  $\mathbf{X}^{(N)} + \mathbf{v}_\tau$ .

The other component of event  $\tau$  is the rate function  $r_\tau^{(N)} : \mathcal{S}^{(N)} \rightarrow \mathbb{R}_{\geq 0}$ , which depends on the current state of the system, and specifies the speed of the corresponding transition. It is assumed to be equal to zero if there are not enough agents available to perform a  $\tau$  transition.

Thus, the population model is  $\mathcal{X}^{(N)} = (\mathbf{X}^{(N)}, \mathcal{T}^{(N)}, \mathbf{x}_0^{(N)})$ , where  $\mathbf{x}_0^{(N)}$  is the initial state. Given such a model, it is straightforward to construct the CTMC  $\mathbf{X}^{(N)}(t)$  associated with it, exhibiting its infinitesimal generator matrix. First, its state space is  $\mathcal{S}^{(N)}$ , while its infinitesimal generator matrix  $Q^{(N)}$  is the  $|\mathcal{S}^{(N)}| \times |\mathcal{S}^{(N)}|$  matrix defined by  $q_{\mathbf{x}, \mathbf{x}'} = \sum \{r_\tau(\mathbf{x}) \mid \tau \in \mathcal{T}, \mathbf{x}' = \mathbf{x} + \mathbf{v}_\tau\}$ .

## 2.2 Running Example: a Bike Sharing system

We consider a bike sharing system with  $N$  members,  $B < N$  bikes, and  $S > B$  bike slots. For simplicity, we assume that the membership of the system is stable within the time period considered, so that the number of members is constant and the closed world assumption holds. Initially members are *Absent* from the system, i.e. engaged in non-transport related activities in their lives. At some point they seek to make a journey through acquiring a bike from the system. A bike may or may not be available to them, so they may successfully make the transition to being a *Biker* (`acq`) or they may have to search more persistently (e.g. visit a bike station in a different location) to find a bike (`fail_acq`); then they enter the state *SeekB*. If on a second attempt the member does not acquire a bike (`fail_acq2`) they become

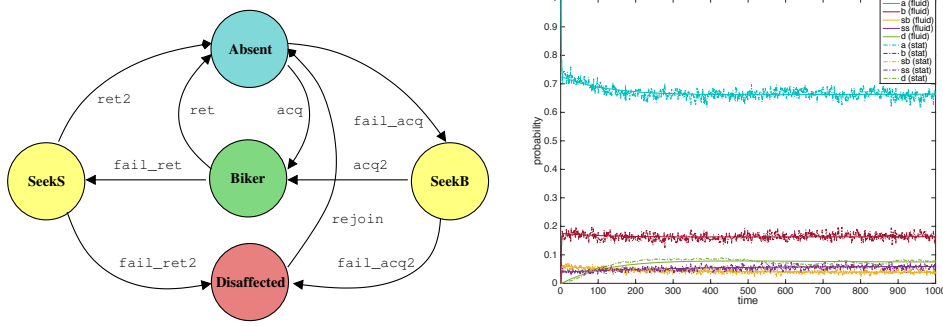


Figure 1: States and transitions of a single member in the bike sharing system (left). Comparison of the solution of the Kolmogorov equation for  $z(t)$  with the statistical estimation (5000 runs) of the state probabilities for  $Z^{(N)}(t)$ . Parameters are  $N = 300$ ,  $S = 150$ ,  $B = 100$ ,  $k_{acq} = 0.25$ ,  $k_{acq2} = 2$ ,  $k_{ret} = 1$ ,  $k_{ret2} = 2$ ,  $k_{rej} = 0.005$ ,  $h = 0.05$ ,  $q = 0.1$ ,  $X_a(0) = N$ ,  $X_i(0) = 0$ , for  $i \neq a$ .

*Disaffected* with probability  $q$  and stop using the system. At the end of a journey the Biker seeks to return the bike to a vacant slot in the system. Again this is a probabilistic action with the probability of success depending on the number of bikes currently in use. If the Biker successfully returns the bike (*ret*) they again become *Absent*; alternatively (*fail\_ret*) they enter the state *SeekS* seeking out a vacant slot into which the bike can be returned (*ret2*). Again, persistent failure (*fail\_ret2*) to achieve their goal can make the member *Disaffected*, with a probability  $q$ . *Disaffected* users, after some amount of time, can rejoin the system, entering into the *Absent* state (*rejoin*). For convenience below we refer to the states as  $a, b, sb, ss$  and  $d$ , with the obvious mapping.

To describe the system in the modelling language we need to specify the collective variables, which in this case are five:  $X_a$  for absent members,  $X_b$  for members using a bike,  $X_{sb}$  for members seeking a bike,  $X_{ss}$  for members seeking a slot and  $X_d$  for disaffected members. Furthermore we need nine transitions or events whose rate and rule sets are described below. Note that the factor  $h$ ,  $0 < h \leq 1$  allows some probabilities to be adjusted according to the information available (see below).

- *acq*:  $R_{acq} = \{a \longrightarrow b\}$   $r_{acq}^{(N)} = k_{acq} p_{bike} X_a$
- *fail\_acq*:  $R_{fail\_acq} = \{a \longrightarrow sb\}$   $r_{fail\_acq}^{(N)} = k_{acq} (1 - p_{bike}) X_a$
- *acq2*:  $R_{acq2} = \{sb \longrightarrow b\}$   $r_{acq2}^{(N)} = k_{acq2} p_{bike}^h X_{sb}$
- *fail\_acq2*:  $R_{fail\_acq2} = \{sb \longrightarrow d\}$   $r_{fail\_acq2}^{(N)} = k_{acq2} (1 - p_{bike}^h) X_{sb}$
- *ret*:  $R_{ret} = \{b \longrightarrow a\}$   $r_{ret}^{(N)} = k_{ret} p_{slot} X_b$
- *fail\_ret*:  $R_{fail\_ret} = \{b \longrightarrow ss\}$   $r_{fail\_ret}^{(N)} = k_{ret} (1 - p_{slot}) X_b$
- *ret2*:  $R_{ret2} = \{ss \longrightarrow a\}$   $r_{ret2}^{(N)} = k_{ret2} p_{slot}^h X_{ss}$
- *fail\_ret2*:  $R_{fail\_ret2} = \{ss \longrightarrow d\}$   $r_{fail\_ret2}^{(N)} = k_{ret2} (1 - p_{slot}^h) X_{ss}$
- *rejoin*:  $R_{rejoin} = \{d \longrightarrow a\}$   $r_{rejoin}^{(N)} = k_{rej} X_d$

Each member seeking a bike has a probability of success  $p_{bike} = p_{bike}(X_b, X_{ss})$  that depends on the number of bikes available, i.e.  $p_{bike}(X_b, X_{ss}) = (B - (X_b + X_{ss}))/B$ . Then the total rate at which bikes are acquired is thus  $k_{acq}p_{bike}(X_b, X_{ss})$ , multiplied by the number of members ready to start their journey,  $X_a$ . The probability of acquiring a bike after a first failure, instead, is  $p_{bike}(X_b, X_{ss})^h$ , where  $h$  models the fact that this probability can be increased by the presence of some information in the bike station just visited (like a screen showing the number of free bikes in nearby stations). Similarly, the rate at which a bike is left depends on the probability of finding a free slot, equal to the fraction of available free slots:  $p_{slot} = p_{slot}(X_b, X_{ss}) = (S - (B - (X_b + X_{ss}))) / S$ .

### 2.3 Fluid limits

In this section we will introduce some concepts of fluid approximation and mean field theory. The basic idea is to approximate a CTMC by an Ordinary Differential Equation (ODE), which can be interpreted in two different ways: it can be seen as an approximation of the average of the system (usually a first order approximation, see [9, 34]), or as an approximate description of system trajectories for large populations. We will focus on this second interpretation, which corresponds to a functional version of the law of large numbers: instead of having a sequence of random variables, like the sample mean, converging to a deterministic value, like the true mean, in this case we have a sequence of CTMCs (which can be seen as random trajectories in  $\mathbb{R}^n$ ) for increasing population size, which converge to a deterministic trajectory — the solution of the fluid ODE.

First we formally define the sequence of CTMCs to be considered. Specifically we normalise the population counts by dividing each variable by the total population  $N$ . The so-obtained normalised population variables  $\hat{\mathbf{X}}^{(N)} = \frac{\mathbf{X}^{(N)}}{N}$ , or population densities, will always range between 0 and 1; thus the behaviour for different population sizes can be compared. We also impose an appropriate scaling to update vectors, initial conditions, and rate functions of the normalised models. Let  $\mathcal{X}^{(N)} = (\mathbf{X}^{(N)}, \mathcal{T}^{(N)}, \mathbf{X}_0^{(N)})$  be the non-normalised model with total population  $N$  and  $\hat{\mathcal{X}}^{(N)} = (\hat{\mathbf{X}}^{(N)}, \hat{\mathcal{T}}^{(N)}, \hat{\mathbf{X}}_0^{(N)})$  the corresponding normalised model. We require that:

- initial conditions scale appropriately:  $\hat{\mathbf{X}}_0^{(N)} = \frac{\mathbf{X}_0^{(N)}}{N}$ ;
- for each transition  $(R_\tau, r_\tau^{(N)}(\mathbf{X}))$  of the non-normalised model, let  $\hat{r}_\tau^{(N)}(\hat{\mathbf{X}})$  be the rate function expressed in the normalised variables (obtained from  $r_\tau^{(N)}$  by a change of variables). The corresponding transition in the normalised model is  $(R_\tau, \hat{r}_\tau^{(N)}(\hat{\mathbf{X}}))$ , with update vector equal to  $\frac{1}{N}\mathbf{v}_\tau$ .

We further assume, for each transition  $\tau$ , that there exists a bounded and Lipschitz continuous function  $f_\tau(\hat{\mathbf{X}}) : E \rightarrow \mathbb{R}^n$  on normalised variables (where  $E$  contains all domains of all  $\hat{\mathcal{X}}^{(N)}$ ), independent of  $N$ , such that  $\frac{1}{N}\hat{r}_\tau^{(N)}(\mathbf{x}) \rightarrow f_\tau(\mathbf{x})$  uniformly on  $E$ . In accordance with Subsection 2.1, we will denote the state of the CTMC of the  $N$ -th non-normalised (resp. normalised) model at time  $t$  as  $\mathbf{X}^{(N)}(t)$  (resp.  $\hat{\mathbf{X}}^{(N)}(t)$ ).

#### 2.3.1 Deterministic limit theorem

Consider a sequence of normalised models  $\hat{\mathcal{X}}^{(N)}$  and let  $\mathbf{v}_\tau$  be the (non-normalised) update vectors. The drift  $F^{(N)}(\hat{\mathbf{X}})$  of  $\hat{\mathcal{X}}^{(N)}$ , the mean instantaneous increment of model variables in state  $\hat{\mathbf{X}}$ , is defined as

$$F^{(N)}(\hat{\mathbf{X}}) = \sum_{\tau \in \hat{\mathcal{T}}} \frac{1}{N} \mathbf{v}_\tau \hat{r}_\tau^{(N)}(\hat{\mathbf{X}}). \quad (2.2)$$

Furthermore, let  $f_\tau : E \rightarrow \mathbb{R}^n$ ,  $\tau \in \hat{\mathcal{T}}$  be the limit rate functions of transitions of  $\hat{\mathcal{X}}^{(N)}$ . The *limit drift* of the model  $\hat{\mathcal{X}}^{(N)}$  is therefore

$$F(\hat{\mathbf{X}}) = \sum_{\tau \in \hat{\mathcal{T}}} \mathbf{v}_\tau f_\tau(\hat{\mathbf{X}}), \quad (2.3)$$

and  $F^{(N)}(\mathbf{x}) \rightarrow F(\mathbf{x})$  uniformly, as easily checked. The fluid ODE is

$$\frac{d\mathbf{x}}{dt} = F(\mathbf{x}), \quad \text{with } \mathbf{x}(0) = \mathbf{x}_0 \in S.$$

Given that  $F$  is Lipschitz in  $E$  (since all  $f_\tau$  are), this ODE has a unique solution  $\mathbf{x}(t)$  in  $E$  starting from  $\mathbf{x}_0$ . Then, one can prove the following theorem:

**Theorem 1 (Deterministic approximation [23, 16])** *Let the sequence  $\hat{\mathbf{X}}^{(N)}(t)$  of Markov processes and  $\mathbf{x}(t)$  be defined as above, and assume that there is some point  $\mathbf{x}_0 \in S$  such that  $\hat{\mathbf{X}}^{(N)}(0) \rightarrow \mathbf{x}_0$  in probability. Then, for any finite time horizon  $T < \infty$ , it holds that:*

$$\mathbb{P} \left\{ \sup_{0 \leq t \leq T} \|\hat{\mathbf{X}}^{(N)}(t) - \mathbf{x}(t)\| > \varepsilon \right\} \rightarrow 0.$$

Note that Theorem 1 is defined with respect to a finite time horizon and does not, in general, tell us anything about the behaviour of the system at steady state. However, there are situations in which we can extend the validity of the theorem to the whole time domain, but this extension depends on properties of the phase space of the fluid ODE [6, 13]. More specifically, we need to require that ODEs have a unique, globally attracting, steady state  $\mathbf{x}^*$ , i.e. a point such that  $\lim_{t \rightarrow \infty} \mathbf{x}(t) = \mathbf{x}^*$  independently of  $\mathbf{x}(0)$ . In those cases, we can prove convergence of the steady state behaviour of  $\hat{\mathbf{X}}^{(N)}(t)$  to that of  $\mathbf{x}(t)$ :

**Theorem 2 (Fluid approximation of steady state [6])** *Let the sequence  $\hat{\mathbf{X}}^{(N)}(t)$  and  $\mathbf{x}(t)$  satisfy hypothesis of Theorem 1, and let  $\mathbf{x}(t)$  have a unique, globally attracting equilibrium  $\mathbf{x}^*$ , and  $\hat{\mathbf{X}}^{(N)}(t)$  have a unique steady state measure  $\hat{\mathbf{X}}^{(N)}(\infty)$  for each  $N$ . Then  $\lim_{N \rightarrow \infty} \|\hat{\mathbf{X}}^{(N)}(\infty) - \mathbf{x}^*\| = 0$  in probability.*

### 2.3.2 Fast simulation

If we focus on a single individual when the population size goes to infinity, even if the collective behaviour tends to a deterministic process, the individual agent will still behave randomly. Moreover, the fluid limit theorem implies that the dynamics of a single agent, in the limit, becomes independent of other agents, sensing them only through the collective system state, described by the fluid limit. This asymptotic decoupling allows us to find a simple, time-inhomogenous, Markov chain for the evolution of the single agent, a result often known as *fast simulation* [17, 18]. More formally, consider a single individual  $Y_h^{(N)}(t)$ , which is a (Markov) process on the state space  $S = \{1, \dots, n\}$ , conditional on the global state of the population  $\hat{\mathbf{X}}^{(N)}(t)$ . Denote by  $Q^{(N)}(\mathbf{x})$  the infinitesimal generator matrix of  $Y_h^{(N)}$ , described as a function of the normalised state of the population  $\hat{\mathbf{X}}^{(N)} = \mathbf{x}$ , i.e.

$$\mathbb{P}\{Y_h^{(N)}(t+dt) = j \mid Y_h^{(N)}(t) = i, \hat{\mathbf{X}}^{(N)}(t) = \mathbf{x}\} = q_{i,j}^{(N)}(\mathbf{x})dt.$$

We stress that  $Q^{(N)}(\mathbf{x})$  describes the exact dynamics of  $Y_h^{(N)}$ , conditional on  $\hat{\mathbf{X}}^{(N)}(t)$ , and that this process is *not independent* of  $\hat{\mathbf{X}}^{(N)}(t)$ .<sup>1</sup> The rate matrix  $Q^{(N)}(\mathbf{x})$  can be constructed from the rate functions

<sup>1</sup>In fact, the marginal distribution of  $Y_h^{(N)}(t)$  is not a Markov process. This means that in order to capture its evolution in a Markovian setting, one has to consider the whole Markov chain  $(Y_h^{(N)}(t), \hat{\mathbf{X}}^{(N)}(t))$ , c.f. also [12] for further details.

of global transitions by computing the fraction of the global rate seen by an individual agent that can perform it. To be more precise, let  $r_\tau^{(N)}(\mathbf{X})$  be the rate function of transition  $\tau$ , and suppose  $i \rightarrow j \in R_\tau$  (and each update rule in  $R_\tau$  has multiplicity one). Then, transition  $\tau$  will contribute to the  $ij$ -entry  $q_{ij}^{(N)}(\mathbf{X})$  of the matrix  $Q^{(N)}(\mathbf{X})$  with the term  $g_{\tau,i}^{(N)}(\hat{\mathbf{X}}) = \frac{1}{\hat{x}_i} r_\tau^{(N)}(\mathbf{X}) = \frac{1}{\hat{x}_i} \hat{r}_\tau^{(N)}(\hat{\mathbf{X}})$ , which converges to  $g_{\tau,i}(\hat{\mathbf{X}}) = \frac{1}{\hat{x}_i} f_\tau(\hat{\mathbf{X}})$ . We define  $g_{\tau,i}(\hat{\mathbf{X}})$  to be the function identical to zero if  $\tau$  is not an action enabled in state  $s_i$ . Additional details of this construction (taking multiplicities properly into account) can be found in [11, 12]. From the previous discussion, it follows that the local rate matrix  $Q^{(N)}(\mathbf{x})$  converges uniformly to a rate matrix  $Q(\mathbf{x})$ , in which all rate functions  $\hat{r}_\tau^{(N)}$  are replaced by their limit  $f_\tau$ . We now define two processes which will be used extensively later:

- $Z^{(N)}(t)$ , which is the stochastic process describing the state of a random individual  $Y_h^{(N)}(t)$  in a population of size  $N$ , marginalised with respect to the collective state  $\hat{\mathbf{X}}^{(N)}(t)$ .
- $z(t)$ , which is a time-inhomogeneous CTMC (ICTMC), on the same state space  $S$  of  $Z^{(N)}$ , with time-dependent rate matrix  $Q(\hat{\mathbf{x}}(t))$ , where  $\hat{\mathbf{x}}(t)$  is the solution of the fluid equation.

The following theorem can be proved [17]:

**Theorem 3 (Fast simulation)** *For any  $T < \infty$ ,  $\mathbb{P}\{Z^{(N)}(t) \neq z(t), \text{ for some } t \leq T\} \rightarrow 0$ , as  $N \rightarrow \infty$ .*

This theorem states that, in the limit of an infinite population, each agent will behave independently from all the others, sensing only the mean state of the global system, described by the fluid limit  $\mathbf{x}(t)$ . This *asymptotic decoupling* of the system, which can be generalised to any subset of  $k \geq 1$  agents, is also known in the literature under the name of *propagation of chaos* [5].

## 2.4 Continuous Stochastic Logic

We now turn to the class of properties that we are interested in checking, which will be specified by the time-bounded fragment of Continuous Stochastic Logic (CSL) (we will introduce reward properties in Section 3). The starting point is a generic labelled stochastic process [3, 4], a random process  $Z(t)$ , with state space  $S$  and a labelling function  $L : S \rightarrow 2^{\mathcal{P}}$ , associating with each state  $s \in S$ , a subset of atomic propositions  $L(s) \subset \mathcal{P} = \{a_1, \dots, a_k \dots\}$  true in that state: each atomic proposition  $a_i \in \mathcal{P}$  is true in  $s$  if and only if  $a_i \in L(s)$ . All subsets of paths considered are provably measurable.

A path of  $Z(t)$  is a sequence  $\sigma = s_0 \xrightarrow{t_0} s_1 \xrightarrow{t_1} \dots$ , such that, given that the process is in  $s_i$  at time  $t_\sigma[i] = \sum_{j=0}^i t_j$ , the probability of moving from  $s_i$  to  $s_{i+1}$  is greater than zero. For CTMCs, this condition is equivalent to  $q_{s_i, s_{i+1}}(t_\sigma[i]) > 0$ , where  $Q = (q_{ij})$  is the infinitesimal generator matrix and  $t_\sigma[i]$  the time of the  $i$ -th jump in  $\sigma$ . We denote by  $\sigma @ t$  the state of  $\sigma$  at time  $t$ , with  $\sigma[i]$  the  $i$ -th state of  $\sigma$ .

A time-bounded CSL formula  $\varphi$  is defined by the following syntax:

$$\varphi ::= \text{true} \mid a \mid \varphi_1 \wedge \varphi_2 \mid \neg \varphi \mid P_{\bowtie p}(\psi) \quad \psi ::= X_{[T_1, T_2]} \varphi \mid \varphi_1 U_{[T_1, T_2]} \varphi_2$$

where  $a$  is an atomic proposition,  $p \in [0, 1]$  and  $\bowtie \in \{<, >, \leq, \geq\}$ .  $\varphi$  are known as *state formulae* and  $\psi$  are *path formulae*. The satisfiability relation of  $\varphi$  with respect to a labelled stochastic process  $Z(t)$  is given by the following rules:

- $s, t_0 \models a$  if and only if  $a \in L(s)$ ;
- $s, t_0 \models \neg \varphi$  if and only if  $s, t_0 \not\models \varphi$ ;
- $s, t_0 \models \varphi_1 \wedge \varphi_2$  if and only if  $s, t_0 \models \varphi_1$  and  $s, t_0 \models \varphi_2$ ;
- $s, t_0 \models P_{\bowtie p}(\psi)$  if and only if  $\mathbb{P}\{\sigma \mid \sigma, t_0 \models \psi\} \bowtie p$ .



- $\sigma, t_0 \models X_{[T_1, T_2]} \varphi$  if and only if  $t_\sigma[1] \in [T_1, T_2]$  and  $\sigma[1], t_0 + t_\sigma[1] \models \varphi$ .
- $\sigma, t_0 \models \varphi_1 U_{[T_1, T_2]} \varphi_2$  if and only if  $\exists \bar{t} \in [t_0 + T_1, t_0 + T_2]$  s.t.  $\sigma @ \bar{t}, \bar{t} \models \varphi_2$  and  $\forall t_0 \leq t < \bar{t}, \sigma @ t, t \models \varphi_1$ .

## 2.5 Fluid Model Checking

The basic idea to approximately check individual CSL properties on large population models is to verify them on the time-inhomogenous CTMC (ICTMC)  $z(t)$  rather than on the process  $Z^{(N)}(t)$ . In [12], we proved that this is a consistent operation, in the sense that:

1. time-bounded reachability probabilities computed for  $Z^{(N)}(t)$  converge to those computed for  $z(t)$ .
2. truth-values of (almost all) state formulae  $\varphi$  converge, in the sense that  $\varphi$  is true in  $z(t)$  if and only if it is true on all  $Z^{(N)}(t)$  for  $N$  large enough.

These convergence results hold for almost all formulae: one has to prohibit the use of some thresholds in the probabilistic quantifier  $P_{\bowtie p}(\psi)$  to avoid computability issues and to ensure that the property  $P_{\bowtie p}(\psi)$  is decidable. (Intuitively, we want the path probability of  $\psi$  to differ from the threshold  $p$ .)

We now provide a few more details of the model checking procedure for ICTMC. For simplicity, we consider a non-nested until formula  $P_{\bowtie p}(\psi)$ , with  $\psi = \varphi_1 U_{[0, T]} \varphi_2$ , in which  $\varphi_1$  and  $\varphi_2$  do not contain any probability quantifier. In order to check it, we need to compute the probability of reaching a  $\varphi_2$ -state without passing through a  $\neg\varphi_1$  state. This can be done by a modification of the method of [4]. We start by making  $\neg\varphi_1$  and  $\varphi_2$  states absorbing, and computing the transient probability matrix  $\Pi(0, T)$  at time  $T$  of the modified CTMC<sup>2</sup>, e.g. by solving the forward Kolmogorov equation [27] or using uniformisation for ICTMC [2]. Note that for ICTMC generally  $\Pi(0, T) \neq \Pi(t, t + T)$ , i.e. the reachability probability depends on the initial time at which it is evaluated. Combining the forward and backward Kolmogorov equations, we obtain a differential equation to compute  $\Pi(t, t + T)$  as a function of the initial time  $t$ :

$$\frac{d}{dt} \Pi(t, t + T) = \Pi(t, t + T) Q(t + T) - Q(t) \Pi(t, t + T). \quad (2.4)$$

Once we have (an approximation of) the function  $\Pi(t, t + T)$ , we just need to compare it with the threshold  $p$ , which can be done by computing the zeros of  $(\Pi(t, t + T) - p)$ ,<sup>3</sup> producing a time-dependent boolean signal, i.e. a function associating with each time  $t$  the truth value of  $P_{\bowtie p}(\psi)$  at that time.

For nested formulae, one has to deal with time dependent truth, i.e. the set of  $\varphi_1$  and  $\varphi_2$  states can change as time passes, and one has to carefully take this into account when computing reachability probabilities by solving equation (2.4). In particular, this introduces discontinuities in such probabilities. To see this, consider a state  $s$  which becomes a goal state (i.e. its satisfaction of  $\varphi_2$  changes from false to true) at a certain time  $t^*$ . Then the probability of being in  $s$  at time  $t^*$  has to be added to the reachability probability since trajectories that are in  $s$  suddenly satisfy the path formula  $\psi$ , introducing a discontinuity.

The general approach is based on defining suitable differential equations to compute the quantities of interest and coupling them with the fluid limit (see [11, 12] for details). We will follow a similar strategy to deal with rewards.

## 3 Checking Rewards

In this section we show how to extend the fluid model checking method to compute reward properties for individual agents. We will first define the reward structures and the reward properties we are interested

<sup>2</sup> $\Pi_{s, s'}(t, t')$  is the probability of being in state  $s'$  at time  $t'$ , given that we were in state  $s$  at time  $t$ .

<sup>3</sup>To ensure these are finite, we restrict all rate functions of the population models to (piecewise) real analytic functions [22].

in. Then, we will show how to compute them, and provide convergence results that ensure that the values we will compute will be consistent with the limit results of Section 2.3.

### 3.1 Reward Structures and CSL Reward Properties

The first step is to extend the individual agent model with a reward structure  $rw$ , which will be composed of a *state reward* function  $\rho_s^{rw}$  and a *transition reward* function  $\rho_t^{rw}$ . As customary, the first function gives the non-negative reward of spending one time unit in any state, while the second encodes the non-negative reward for taking a certain transition. More specifically, let  $S$  be the state space of an individual agent, and  $\mathcal{T}$  be the set of transitions of the population model (where we dropped the superscript  $N$  as we only care about the transition per se, not its  $N$ -dependent rate function):

- The state reward is  $\rho_s^{rw} : S \rightarrow \mathbb{R}_{\geq 0}$
- The transition reward is  $\rho_t^{rw} : \mathcal{T} \rightarrow \mathbb{R}_{\geq 0}$

Given a reward structure  $rw$  and a stochastic process  $Z(t)$  on the state space  $S$  of an individual agent, we will consider four kinds of rewards:

- **Cumulative rewards** up to time  $T$ . Given a trajectory  $\sigma = s_0 \xrightarrow{t_0, \alpha_0} s_1 \xrightarrow{t_1, \alpha_1} \dots s_n$  with a time span of  $T$  (hence remaining some time in the final state  $s_n$ ), its cumulative reward  $\rho_c^{rw}(\sigma)$  is defined as

$$\rho_c^{rw}(\sigma) = \sum_{i=0}^{n-1} \rho_s^{rw}(s_i) t_i + \rho_s^{rw}(s_n) (T - \sum_i t_i) + \sum_{i=0}^{n-1} \rho_t^{rw}(\alpha_i).$$

The first two terms represent the state reward accumulated by remaining in states  $s_i$ , while the last term of the sum is the transition reward accumulated by the jumps in  $\sigma$ . The expected cumulative reward  $\rho_c^{rw}(T)$  is just the expectation of  $\rho_c^{rw}(\sigma)$  over all trajectories with time span restricted to  $[0, T]$ . Similarly, we define  $\rho_c^{rw}(T, s)$  as the expectation over all trajectories  $\sigma$  such that  $\sigma_0 = s$  and time span restricted as above.

- **Instantaneous rewards** at time  $T$ .  $\rho_t^{rw}(T, s)$  is defined as the expected value of  $\rho_s^{rw}$  at time  $T$ , conditional on starting from state  $s$ , i.e.

$$\rho_t^{rw}(T, s) = \sum_{s' \in S} \rho_s^{rw}(s') \mathbb{P}\{Z(T) = s' \mid Z(0) = s\}.$$

- **Steady state rewards.**  $\rho_{ss}^{rw}$  is the expected reward at steady state, assuming the process  $Z(t)$  has a unique steady state measure, independent of the initial state.
- **Bounded reachability rewards.** Consider a subset of states  $A \subseteq S$ , then  $\rho_{reach}(A, T, s)$  is the cumulative reward starting from state  $s$ , until we enter an  $A$ -state, within the time-horizon  $T$ . Formally,  $\rho_{reach}(A, T, s)$  is defined as the cumulative reward up to time  $T$  for the modified processes  $Z_A^{(N)}$  and  $z_A$ , in which  $A$  states are made absorbing (by removing outgoing transitions), and for the modified state reward  $\rho_{s|A}$ , defined by  $\rho_{s|A}(s) = \rho_s(s)$  if  $s \notin A$  and 0 if  $s \in A$ .

In the context of this paper, we will express these reward properties by extending the time-bounded CSL fragment with the following reward operators:

$$\varphi ::= \mathbf{R}_{\bowtie T}[rw, C \leq T] \mid \mathbf{R}_{\bowtie T}[rw, I = T] \mid \mathbf{R}_{\bowtie T}[rw, S] \mid \mathbf{R}_{\bowtie T}[rw, \mathbf{F}_{\leq T} \varphi],$$

where we assume that  $\varphi$  is a CSL formula not containing any temporal operator.<sup>4</sup> The semantics of these operators is defined for a stochastic process  $Z(t)$  over  $S$  and a reward structure  $rw$  (which is made explicit in the formula), by comparing the suitable expected reward with the threshold  $r \geq 0$ :

- $s, t \models \mathbf{R}_{\bowtie r}[rw, C \leq T]$  if and only if  $\rho_c^{rw}(T, s) \bowtie r$ ;
- $s, t \models \mathbf{R}_{\bowtie r}[rw, I = T]$  if and only if  $\rho_I^{rw}(T, s) \bowtie r$ ;
- $s, t \models \mathbf{R}_{\bowtie r}[rw, S]$  if and only if  $\rho_{ss}^{rw} \bowtie r$ ;
- $s, t \models \mathbf{R}_{\bowtie r}[rw, \mathbf{F}_{\leq T} \varphi]$  if and only if  $\rho_{reach}^{rw}(S_\varphi, T, s) \bowtie r$ , where  $S_\varphi = \{s \in S \mid s \models \varphi\}$ .

**Example** Consider again the bike sharing model introduced in Section 2.2. First, we want to model the fact that there is a charge for using the bike sharing system which depends on the time for which a bike is in use. This is represented by a state reward equal to  $\kappa$ , which is associated with the states *Biker* and *SeekS*. Hence, the corresponding reward structure is given by  $\rho_s^{cost}(Biker) = \rho_s^{cost}(SeekS) = \kappa$ , and  $\rho_s^{cost}(s') = 0$  for all other states. The transition reward is set identically equal to zero. Thus the total expenditure of the member up to a given time is given by the cumulative reward associated with the *cost* reward structure, and the fact that this cost is less than  $r$  can be formally expressed by the CSL property  $\mathbf{R}_{\leq r}[cost, C \leq T]$ . We can also compute the cost until an agent becomes dissatisfied, as  $\mathbf{R}_{=?}[cost, \mathbf{F}_{\leq T} at_d]$ . Here the notation  $=?$  is taken from PRISM [1] and, as customary, denotes that we are interested in the value of the reward, rather than in comparing its value with a threshold.

Dissatisfaction with the service, instead, is incurred whenever the member is not able to obtain a bike or return a bike on the first attempt. This is captured by the reward structure *diss*, with a transition reward equal to  $\delta$  for the actions *fail\_acq* and *fail\_ret* and to  $10 \times \delta$  for the actions *fail\_acq2* and *fail\_ret2*. Here the instantaneous value of the reward will give the expected level of dissatisfaction of the member. The requirement that this is below the value  $r$  is encoded by the formula  $\mathbf{R}_{\leq r}[diss, I = T]$ .

## 3.2 Fluid Approximation of Individual Rewards

We now describe how to approximate the rewards of an individual agent in a large model. The idea is simple: we will just replace the agent  $Z^{(N)}(t)$  by the agent  $z(t)$  operating in the mean field environment. In the following, we will sketch the algorithms to compute such rewards, and prove that the reward of  $Z^{(N)}(t)$  will converge to the one of  $z(t)$ , as  $N$  goes to infinity. This will prove the consistency of the approximation.

### 3.2.1 Instantaneous rewards

We will first consider the instantaneous reward  $\rho_I$ , omitting the reward structure from the notation, which is assumed to be fixed. To further fix the notation, we will call  $\rho_I(T)$  the instantaneous reward at time  $T$  for  $z(t)$ , and  $\rho_I^{(N)}(T)$  the same reward for  $Z^{(N)}(t)$ .

**Algorithm.** Computing the instantaneous reward  $\rho_I^{(N)}(T)$  for  $Z^{(N)}(t)$  requires knowledge of the transient distribution of  $Z^{(N)}(T)$  at time  $T$ . This can be obtained from the Markov Chain  $(Z^{(N)}, \mathbf{X}^{(N)})$ , in which the individual agent is tracked in the population model, and computing the transient distribution of this process, then marginalising it to  $Z^{(N)}(T)$ . Solving the larger process is necessary, because  $Z^{(N)}(T)$

<sup>4</sup>This is for simplicity, as the general case requires us to work with the time-dependent truth of  $\varphi$ , which makes the algorithm more involved by introducing discontinuities in the reward, c.f. also the discussion in Section 2.3.1.

is not Markovian: its evolution depends on  $\mathbf{X}^{(N)}$ . It follows that, for large  $N$ , the cost of this operation becomes prohibitive. However, we will approximate  $\rho_I^{(N)}(T)$  by  $\rho_I(T)$ , the reward of the limit model  $z(t)$ . This is much easier to compute: we just need to compute its transient distribution at time  $T$ , which can be done either by uniformisation for ICTMC, or by solving directly the Kolmogorov equations, coupled with the fluid equations whose solution defines the rates of  $z(t)$ . More specifically, we need to solve the following initial value problem:

$$\frac{d\mathbf{x}}{dt} = F(\mathbf{x}), \quad \frac{dP}{dt} = PQ(\mathbf{x}), \quad \text{with } \mathbf{x}(0) = \mathbf{x}_0, \quad P(0) = \delta_s, \quad s \in S. \quad (3.5)$$

where  $P(t)$  is the probability distribution over the state space of the agent, with initial distribution  $\delta_s$ . Then the instantaneous reward is computed as

$$\rho_I(T) = \sum_s \rho_s(s) P(s).$$

**Convergence.** The fact that  $\lim_{N \rightarrow \infty} \rho_I^{(N)}(T) = \rho_I(T)$  is essentially a corollary of Theorem 3. For completeness, we sketch the proof. By standard arguments in probability theory [7], it follows from Theorem 3 that  $\mathbf{Z}^{(N)} \Rightarrow z$  weakly, when  $\mathbf{Z}^{(N)}$  and  $z$  are seen as random variables on the space of trajectories  $\mathcal{D}_S$ , i.e. of  $S$ -valued cadlag functions<sup>5</sup>. We can define a functional  $R_{I=T}$  on  $\mathcal{D}_S$  that, given a trajectory  $\sigma : [0, \infty) \rightarrow S$ , returns the value of the state reward at time  $T$ , i.e.  $R_{I=T}(\sigma) = \rho_s(\sigma(T))$ . Obviously,  $R_{I=T}$  is bounded by  $\max_{s \in S} \rho_s(s)$ , and it is furthermore continuous on all trajectories that do not jump at time  $T$ . Since  $\mathbf{Z}^{(N)} \Rightarrow z$  weakly and this set of trajectories accumulates probability one (the probability of jumping at  $T$  is zero), by the Portmanteau theorem<sup>6</sup>, it holds that

$$\rho_I^{(N)}(T) = \mathbb{E}[R_{I=T}(\mathbf{Z}^{(N)})] \xrightarrow{N \rightarrow \infty} \mathbb{E}[R_{I=T}(z)] = \rho_I(T).$$

### 3.2.2 Cumulative rewards

We now turn our attention to the cumulative reward. We will show how to compute  $\rho_c(T, s)$ , the cumulative reward for  $z(t)$ , and prove that  $\rho_c^{(N)}(T, s)$ , the cumulative reward for  $\mathbf{Z}^{(N)}(t)$ , converges to  $\rho_c(T, s)$  as  $N$  diverges.

**Algorithm.** The computation of  $\rho_c(T, s)$  can be done by augmenting the forward Kolmogorov equations (3.5) for  $z(t)$  with an additional equation for the cumulative reward. It is easy to see that we can express the cumulative reward for the trajectory of  $z(t)$  as an integral, by introducing further random variables  $c_\alpha(t)$  counting how many times the process jumped by taking transition  $\alpha$ :

$$\rho_c(T, s) = \mathbb{E} \left[ \int_0^T \rho_s(z(t)) dt + \sum_\alpha \rho_t(\alpha) c_\alpha(T) \right] = \int_0^T \mathbb{E}[\rho_s(z(t))] dt + \sum_\alpha \rho_t(\alpha) \mathbb{E}[c_\alpha(T)].$$

<sup>5</sup>  $\mathcal{D}_S$  is the space of functions  $\sigma : [0, \infty) \rightarrow S$  that have at most a countable number of discontinuous jumps.  $\mathcal{D}_S$  can be turned into a metric space by the Skorokhod metric, cf [8].

<sup>6</sup> The Portmanteau theorem [8] states that for weakly convergent measures  $\mu_n \Rightarrow \mu$ , the expectation  $\int f d\mu_n$  with respect to  $\mu_n$  will converge to the expectation  $\int f d\mu$  with respect to  $\mu$  for all  $\mu$ -almost surely bounded continuous functions  $f$ , i.e. continuous on a measurable subset  $A \subset \mathcal{D}_S$  such that  $\mu(A) = 1$ .

Differentiating this equation with respect to  $T$ , we get the following ODE for  $\rho_c(T, s)$  (with  $\rho_c(0, s) = 0$ ):

$$\frac{d}{dT}\rho_c(T, s) = \mathbb{E}[\rho_s(z(T))] + \sum_{\alpha} \rho_t(\alpha) \frac{d}{dT}\mathbb{E}[c_{\alpha}(T)] = \sum_{s'} \rho_s(s')P_{ss'}(T) + \sum_{\alpha} \rho_t(\alpha) \sum_{s' \in S} g_{\alpha, s'}(\mathbf{x}(T))P_{ss'}(T) \quad (3.6)$$

Here  $P_{ss'}(T)$  is the probability of being in  $s'$  at time  $T$  given that the process started in state  $s$ , and  $\frac{d}{dT}\mathbb{E}[c_{\alpha}(T)] = \sum_{s' \in S} g_{\alpha, s'}(\mathbf{x}(T))P_{ss'}(T)$ . Hence, the cumulative reward  $\rho_c(T, s)$  can be computed by solving the initial value problem obtained by combining equations (3.5) and (3.6) up to time  $T$ .

**Convergence.** We will show here that the cumulative reward  $\rho_c^{(N)}(T, s)$  for process  $Z^{(N)}(t)$  converges to the reward  $\rho_c(T, s)$  for process  $z(t)$ . As in the previous subsection, we will first define a functional  $R_{c \leq T}(\sigma)$  on the space of trajectories  $\mathcal{D}_S$ , associating the cumulative reward to each trajectory  $\sigma$ , taking its expectation with respect to the measures on  $\mathcal{D}_S$  induced by  $Z^{(N)}$  and  $z$ , respectively. Differently from the previous section, however, we cannot rely on the Portmanteau theorem to conclude, because the functional  $R_{c \leq T}(\sigma)$  is not bounded, as the number of jumps in  $\sigma$  up to time  $T$  is unbounded, and so is the functional  $R_{c \leq T}$ , due to the term that accumulates transition rewards.

To circumvent this problem, we can reason as follows:

1. The entries of the matrix  $Q^{(N)}(\mathbf{x})$  and of  $Q(\mathbf{x})$ , defining the single agent behaviour, are bounded functions: they are continuous functions and, due to the conservation of the population,  $\mathbf{x}$  belongs to a compact subset of  $\mathbb{R}^n$ .<sup>7</sup> Let  $M > 0$  be an upper bound of the transition rates for both the limit process  $Q(\mathbf{x})$  and  $Q^{(N)}(\mathbf{x})$ , for all  $N \geq N_0$ , for some  $N_0$ .
2. The number of jumps of  $Z^{(N)}$  and of  $z$  is then stochastically bounded by a Poisson process with rate  $M$ .
3. The expected reward due to transitions, assuming  $\rho_t(s) \leq K$  for all  $s$ , is then bounded by  $KMT$ :

$$\sum_m m K e^{-MT} \frac{(MT)^m}{m!} \leq KMT \sum_m e^{-MT} \frac{(MT)^{m-1}}{(m-1)!} \leq KMT.$$

Now, let us explicitly split the contributions of state and transition rewards in  $R_{c \leq T}$ :  $R_{c \leq T}(\sigma) = R_{c \leq T}^{\text{state}}(\sigma) + R_{c \leq T}^{\text{trans}}(\sigma)$ . Furthermore, let  $k$  be a number of jumps such that the probability of firing transitions more than  $k$  times is less than  $\frac{\varepsilon}{2KMT}$  for some  $\varepsilon$ , and let  $\Omega_{\leq k}$  be the event that there have been  $k$  jumps or less. Then

$$\begin{aligned} |\mathbb{E}[R_{c \leq T}(Z^{(N)})] - \mathbb{E}[R_{c \leq T}(z)]| &\leq |\mathbb{E}[R_{c \leq T}^{\text{state}}(Z^{(N)})] - \mathbb{E}[R_{c \leq T}^{\text{state}}(z)]| + |\mathbb{E}[R_{c \leq T}^{\text{trans}}(Z^{(N)})] - \mathbb{E}[R_{c \leq T}^{\text{trans}}(z)]| \\ &\leq |\mathbb{E}[R_{c \leq T}^{\text{state}}(Z^{(N)})] - R_{c \leq T}^{\text{state}}(z)| + |\mathbb{E}[R_{c \leq T}^{\text{trans}}(Z^{(N)}) - R_{c \leq T}^{\text{trans}}(z) \mid \Omega_{\leq k}]| \mathbb{P}\{\Omega_{\leq k}\} \\ &\quad + (\mathbb{E}[|R_{c \leq T}^{\text{trans}}(Z^{(N)})| \mid \Omega_{> k}] - \mathbb{E}[|R_{c \leq T}^{\text{trans}}(z)| \mid \Omega_{> k}]) \mathbb{P}\{\Omega_{> k}\} \\ &\leq |\mathbb{E}[R_{c \leq T}^{\text{state}}(Z^{(N)})] - R_{c \leq T}^{\text{state}}(z)| + |\mathbb{E}[R_{c \leq T}^{\text{trans}}(Z^{(N)}) - R_{c \leq T}^{\text{trans}}(z) \mid \Omega_{\leq k}]| + \varepsilon. \end{aligned}$$

Now, in the last inequality, the Portmanteau theorem [8] implies that the first two terms on the right go to zero as  $N$  diverges, as the cumulative state reward functional, and the cumulative transition reward

<sup>7</sup>A similar but slightly more involved argument would apply also to the case of non-conserved populations. In this case, in fact, one can rely on the fact that the solution of the fluid equation in  $[0, T]$  is bounded, say by  $M > 0$ , and then invoke the fluid limit theorem to show that, almost surely, the PCTMC is bounded by  $M + \varepsilon$ , for a fixed  $\varepsilon > 0$  and  $N$  large enough. Then the boundedness of  $Q^{(N)}(\mathbf{x})$  and of  $Q(\mathbf{x})$  follows by continuity of rates.

functional conditional on  $k$  jumps or less, are both almost surely continuous and bounded. It follows that  $\lim_{N \rightarrow \infty} |\mathbb{E}[R_{c \leq T}(Z^{(N)})] - \mathbb{E}[R_{c \leq T}(z)]| \leq \varepsilon$  for each  $\varepsilon > 0$ , implying that  $\lim_{N \rightarrow \infty} |\mathbb{E}[R_{c \leq T}(Z^{(N)})] - \mathbb{E}[R_{c \leq T}(z)]| = 0$ .

### 3.2.3 Steady-State Rewards

Considering the steady state reward  $\rho_{ss}$ , we start by making the assumption that we are in the hypothesis of Theorem 2, i.e. that the fluid convergence theorem can be extended to the steady state behaviour. It then follows that the behaviour at steady state of an individual agent converges also to the steady state behaviour of  $z(t)$ , which itself can be found by computing the invariant measure  $\pi^*(s)$  of the transition matrix  $Q(\mathbf{x}^\infty)$  evaluated at the steady state value  $\mathbf{x}^\infty$  of the fluid limit.

Given  $\pi^*$ , the reward  $\rho_{ss}$  for  $z(t)$  is easily computed as  $\rho_{ss} = \sum_s \rho_s(s) \pi^*(s)$ . By the weak convergence of steady state measures for individual agents (an easy corollary of Theorem 2) it then follows that  $\rho_{ss}^{(N)} \rightarrow \rho_{ss}$ .

### 3.2.4 Time-bounded Reachability Rewards

Finally, the algorithm to compute reachability rewards  $\rho_{reach}(S_\phi, T, s)$  for  $Z(t)$  starts by constructing the modified process  $z_A(t)$  in which  $A$ -states are made absorbing, and by modifying the state reward. Then it applies to this new process the algorithm for cumulative rewards. Convergence follows from the result for cumulative rewards, by replacing  $Z^{(N)}$  with  $Z_A^{(N)}$  and  $z$  by  $z_A$ , and by invoking the fact that the modified process  $Z_A^{(N)}$  converges weakly to the modified process  $z_A$ .

### 3.2.5 Running Example

Finally we provide some experimental evidence of the goodness of the fluid approximation for the bike sharing system example discussed in Section 2.2. Here we consider the CSL reward properties shown in the table below. Figure 2 shows the value of the fluid rewards compared to a statistical estimation of the rewards (1000 runs) in the stochastic model as a function of  $T$ , for  $T \in [0, 1000]$ , for  $\Phi_1$  and  $\Phi_2$ , and  $T \in [0, 10000]$  for  $\Phi_3$ . The choice of the different time bound for  $\Phi_3$  is determined by the fact that this is a reachability property, requiring a modification of the underlying CTMC resulting in a longer time to stabilise. In the table below, we report for the same properties the maximum and average error obtained (parameters as in Figure 1), the relative error at the final time, and the computational cost of the fluid and statistical computations. Results were obtained on a standard laptop, implementing both models in the Java tool SimHyA [10], developed by one of the authors.

| property  | max error | mean error | rel err at $T = 1000$ | cost (stat) | cost (fluid) |
|---|-----------|------------|-----------------------|-------------|--------------|
| $\Phi_1 := \mathbf{R}_{=?}[cost, C \leq T]$                 | 2.01      | 1.05       | 0.004                 | 679.54 sec  | 0.16 sec     |
| $\Phi_2 := \mathbf{R}_{=?}[diss, I = T]$                    | 1.16      | 0.51       | 0.005                 | 676.77 sec  | 0.12 sec     |
| $\Phi_3 := \mathbf{R}_{=?}[cost, \mathbf{F}_{\leq T} at_d]$ | 13.22     | 8.85       | 0.023                 | 6885.03 sec | 1.10 sec     |

As we can see, the results are quite accurate even for a small population of  $N = 300$ , but the computational cost is 3 orders of magnitude smaller (and independent of  $N$  for the fluid case). The computational time for  $\Phi_3$  is larger both in the fluid and the statistical estimate due to the larger time bound considered in the property. As for the quality of the approximation, the performance of the method is worse for  $\Phi_3$  likely because  $\Phi_3$  is a reachability reward property, hence it is subject to two sources of errors: the approximation of the reachability probability and that of the cumulative reward.

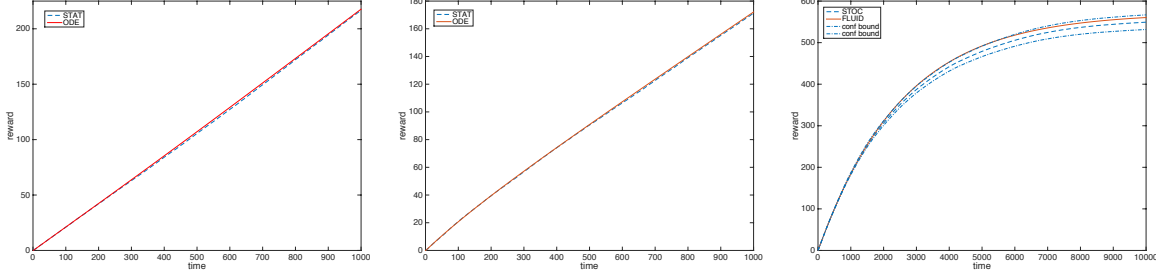


Figure 2: Comparison of reward properties  $\Phi_1$  (left),  $\Phi_2$  (middle), and  $\Phi_3$  (right), computed by the fluid approximation (solid red line) and by stochastic simulation (1000 runs, dashed blue line). In the first two plots, lines almost overlap, while in the third one, the fluid approximation falls within the 95% confidence bounds of the simulation estimate.

## 4 Conclusion

In this paper we extended the fluid model checking framework [11] to deal with reward properties of individual agents. We presented the algorithms and the convergence results, and we also discussed a bike sharing example. This represents a further step to be able to tackle the full set of CSL properties in a consistent way.

The use of time-bounded reachability reward operator is unusual, as this class of rewards is usually specified in a time unbounded sense. Doing this for the fluid approximation, however, introduces an additional challenge: we do not know if the time-unbounded reachability probabilities, and hence the time unbounded reachability rewards, will converge or not, even if the conditions of Theorem 2 are satisfied. We conjecture that this is the case, and that such quantities can be obtained from the corresponding time-bounded one as time goes to infinity, i.e.  $\rho_{reach}(S_\varphi, \infty, s) = \lim_{T \rightarrow \infty} \rho_{reach}(S_\varphi, T, s)$ . Extending the framework to deal with time-unbounded temporal and reward properties, remains as work to be done. Additional improvements of the framework may be obtained by coupling individual agents with a second or higher order approximation of the environment process. We plan to explore these directions in the future.

**Acknowledgements.** Work partially supported by EU-FET project QUANTICOL (nr. 600708), FRA-UniTS, and the German Research Council (DFG) as part of the Cluster of Excellence on Multimodal Computing and Interaction (Saarland University) and Transregional Collaborative Research Center SFB/TR 14 AVACS.

## References

- [1] *PRISM Home Page*. <http://www.prismmodelchecker.org>.
- [2] M. Arns, P. Buchholz & A. Panchenko (2010): *On the Numerical Analysis of Inhomogeneous Continuous-Time Markov Chains*. *INFORMS Journal on Computing* 22(3), pp. 416–432.
- [3] A. Aziz, V. Singhal, F. Balarin, R. Brayton & A. Sangiovanni-Vincentelli (1996): *Verifying Continuous Time Markov Chains*. In: *Proceedings of CAV96*.

- [4] C. Baier, B.R. Haverkort, H. Hermanns & J.P. Katoen (2003): *Model-Checking Algorithms for Continuous-Time Markov Chains*. *IEEE Trans. Software Eng.* 29(6), pp. 524–541, doi:10.1109/TSE.2003.1205180.
- [5] M. Benaïm & J. Le Boudec (2008): *A Class of Mean Field Interaction Models for Computer and Communication Systems*. *Performance Evaluation*.
- [6] M. Benaïm & J.Y. Le Boudec (2011): *On Mean Field Convergence and Stationary Regime*. CoRR abs/1111.5710. Available at <http://arxiv.org/abs/1111.5710>.
- [7] P. Billingsley (1979): *Probability and Measure*. John Wiley and Sons.
- [8] P. Billingsley (1999): *Convergence of Probability Measures, 2nd Edition*. Wiley.
- [9] L. Bortolussi (2008): *On the Approximation of Stochastic Concurrent Constraint Programming by Master Equation*. In: *Proceedings of QAPL*, pp. 163–180, doi:10.1016/j.entcs.2008.11.025.
- [10] L. Bortolussi, V. Galpin & J. Hillston (2012): *Hybrid performance modelling of opportunistic networks*. In: *Proceedings of QAPL 2012, EPTCS 85*, p. 106–121, doi:10.4204/EPTCS.85.8.
- [11] L. Bortolussi & J. Hillston (2012): *Fluid Model Checking*. In: *Proceedings of CONCUR 2012*.
- [12] L. Bortolussi & J. Hillston (2015): *Model Checking Single Agent Behaviour by Fluid Approximation*. *Information and Computation*. To appear.
- [13] L. Bortolussi, J. Hillston, D. Latella & M. Massink (2013): *Continuous Approximation of Collective Systems Behaviour: a Tutorial*. *Perf. Eval.* 70(5), pp. 317–349.
- [14] L. Bortolussi & R. Lanciani (2013): *Model Checking Markov Population Models by Central Limit Approximation*. In: *Proceedings of QEST*, pp. 123–138.
- [15] E. Clarke, A. Peled & A. Grunberg (1999): *Model Checking*. MIT press.
- [16] R.W.R. Darling (2002): *Fluid Limits of Pure Jump Markov Processes: A Practical Guide*. [arXiv.org](http://arxiv.org).
- [17] R.W.R. Darling & J.R. Norris (2008): *Differential equation approximations for Markov chains*. *Probability Surveys* 5.
- [18] N. Gast & B. Gaujal (2010): *A mean field model of work stealing in large-scale systems*. In: *Proceedings of ACM SIGMETRICS 2010*, pp. 13–24. Available at <http://doi.acm.org/10.1145/1811039.1811042>.
- [19] R. A. Hayden, J. T. Bradley & A. Clark (2013): *Performance Specification and Evaluation with Unified Stochastic Probes and Fluid Analysis*. *IEEE Trans. Software Eng.* 39(1), pp. 97–118.
- [20] R.A. Howard (2007): *Dynamic Probabilistic Systems, Volume II*. Dover.
- [21] A. Kolesnichenko, A. Remke, P.T. de Boer & B.R. Haverkort (2011): *Comparison of the Mean-Field Approach and Simulation in a Peer-to-Peer Botnet Case Study*. In: *Proceedings of EPEW*, pp. 133–147. Available at 10.1007/978-3-642-24749-1\_11.
- [22] S. Krantz & P.R. Harold (2002): *A Primer of Real Analytic Functions (Second ed.)*. Birkhäuser.
- [23] T. G. Kurtz (1970): *Solutions of Ordinary Differential Equations as Limits of Pure Jump Markov Processes*. *Journal of Applied Probability* 7, pp. 49–58.
- [24] D. Latella, M. Loretì & M. Massink (2013): *On-the-fly Fast Mean-Field Model-Checking*. In: *Proceedings of TGC*, pp. 297–314, doi:10.1007/978-3-319-05119-2\_17.
- [25] J.-Y. Le Boudec (2010): *Performance Evaluation of Computer and Communication Systems*. EPFL Press, Lausanne.
- [26] M. Massink, D. Latella, A. Bracciali, M. Harrison & J. Hillston (2012): *Scalable context-dependent analysis of emergency egress models*. *Formal Aspects of Computing*, pp. 1–36, doi:10.1007/s00165-011-0188-1.
- [27] J. R. Norris (1997): *Markov Chains*. Cambridge University Press.
- [28] H. Qian & E.L. Elson (2002): *Single-molecule enzymology: stochastic Michaelis-Menten kinetics*. *Biophysical Chemistry* 101, p. 565–576.
- [29] J. Rutten, M. Kwiatkowska, G. Norman & D. Parker (2004): *Mathematical Techniques for Analyzing Concurrent and Probabilistic Systems*. CRM Monograph Series 23, American Mathematical Society.



- [30] A. Stefanek, R. A. Hayden, M. Mac Gonagle & J. T. Bradley (2012): *Mean-Field Analysis of Markov Models with Reward Feedback*. In: *Proceedings of ASMTA 2012*, pp. 193–211, doi:10.1007/978-3-642-30782-9\_14.
- [31] D.T.J. Sumpter (2000): *From Bee to Society: An Agent-based Investigation of Honey Bee Colonies*. Ph.D. thesis, University of Manchester.
- [32] Z. Szallasi, J. Stelling & V. Periwal, editors (2012): *System Modeling in Cellular Biology, From Concepts to Nuts and Bolts*. MIT Press.
- [33] M. Tribastone, J. Ding, S. Gilmore & J. Hillston (2012): *Fluid Rewards for a Stochastic Process Algebra*. *IEEE Trans. Software Eng.* 38(4), pp. 861–874, doi:10.1109/TSE.2011.81.
- [34] N. G. Van Kampen (1992): *Stochastic Processes in Physics and Chemistry*. Elsevier.